

Montclair State University

Montclair State University Digital Commons

Department of Linguistics Faculty Scholarship
and Creative Works

Department of Linguistics

3-1-2000

Types of morphemes and their implications for second language morpheme acquisition

Longxing Wei

Montclair State University, weil@mail.montclair.edu

Follow this and additional works at: <https://digitalcommons.montclair.edu/linguistics-facpubs>

 Part of the [Linguistics Commons](#)

MSU Digital Commons Citation

Wei, Longxing, "Types of morphemes and their implications for second language morpheme acquisition" (2000). *Department of Linguistics Faculty Scholarship and Creative Works*. 39.
<https://digitalcommons.montclair.edu/linguistics-facpubs/39>

This Article is brought to you for free and open access by the Department of Linguistics at Montclair State University Digital Commons. It has been accepted for inclusion in Department of Linguistics Faculty Scholarship and Creative Works by an authorized administrator of Montclair State University Digital Commons. For more information, please contact digitalcommons@montclair.edu.

Types of morphemes and their implications for second language morpheme acquisition

Longxing Wei

Montclair State University

Abstract

This paper explains observed morpheme accuracy orders on the basis of a model of morpheme classification, the 4-M model proposed by Myers-Scotton and Jake (2000). It argues that the adult second language morpheme acquisition order is determined by how morphemes are projected from the mental lexicon. Four types of morphemes are identified: content morphemes, early system morphemes, and two types of late system morphemes. Early system morphemes are indirectly elected at the same time that content morphemes are directly elected by the speaker's intentions. Late system morphemes are activated later in the production process as required by the grammatical frame of the target language. This paper claims that there is variation within individual lexical categories and that the distinction between particular types of morphemes is not a lexical category-defining feature. That is, the classification of morphemes is based on how morphemes are activated. Interlanguage data from early adult Chinese and Japanese learners of English as a second language indicate an implicational hierarchy of morpheme acquisition: content morphemes are acquired before any system morphemes, and early system morphemes are acquired before late system morphemes. Reported in this paper are the learners' production of English determiners and pronouns relevant to testing the categorization of morphemes as specified in the 4-M model. The accuracy/frequency count of the learners' acquisition of the types of morphemes provides statistical evidence for the 4-M model and the hypotheses of the study.

Key words

4-M model

morpheme acquisition

second language acquisition

1. Introduction

This paper proposes that interlanguage (IL) constructions are driven by different types of morphemes. It argues that the reason why morphemes are not acquired at the same rate is that they are projected differently from the mental lexicon. It claims that early IL forms are the consequences of the learner's incomplete acquisition of the abstract lexical entries of the target language. In support of the 4-M model (Myers-Scotton & Jake, 2000), this study identifies how four types of morphemes—content morphemes and three types of system morphemes—play different roles in adult second language morpheme acquisition. Based on IL data from early adult Chinese and Japanese learners of English as a second language, it clarifies and motivates the distinct roles played by different types of morphemes in building IL constructions.

Early studies in morpheme acquisition order of inflectional elements, such as plural -s and 3rd person -s, and grammatical functors, such as article and copula, are descriptive in

Address for correspondence

nature and offer no explanations for the “natural order” of second language sequence (Bailey, Madden, & Krashen, 1974; Dulay & Burt, 1972, 1973, 1974). Dulay, Burt and Krashen state, “Researchers have finally discovered the major reason behind such apparently intractable errors: the third person *-s* and *has* appear relatively late in the order in which learners naturally acquire linguistic structures” (1982, pp. 200–201). But the order is not the reason behind errors, and the order itself is not an explanatory generalization about errors. The proposals in this paper provide a formal theoretical model of lexical structure at some abstract level that formalizes concepts for the characterization of morphemes and functors. This study is based on the assumption that the theoretical constructs of the distinction between content and system morpheme and the distinction between conceptually activated and nonconceptually activated lexemes best capture the generalizations of morpheme accuracy/frequency acquisition.

This paper has two goals. First, on the descriptive level, it characterizes morphemes in terms of their entries in the mental lexicon (i.e., lemmas). Evidence indicates that morpheme acquisition order would not be explained without characterization of morphemes themselves, and such a characterization depends on the formalization of a connection between the underlying abstract lexical entries in the mental lexicon and surface realizations. Second, the study goes beyond describing the nature of different types of morphemes to investigating the relative accessibility of morphemes in second language production. In so doing, it provides explanations and makes predictions for the morpheme acquisition order. The main focus of this paper is to argue that the findings from a project investigating adult second language morpheme acquisition order and its constraints on IL development provide evidence that not all functional elements are equal. The 4-M model captures these distinctions in claiming that not all such elements are accessed in the same way. Differences in degree of accuracy for these elements in IL support this claim. The following morphemes from the larger study are discussed here: under the category Determiner, definite articles, indefinite articles, possessives, and demonstratives; under the category Pronoun, personal pronouns, demonstrative pronouns, and the dummy pronouns *it* and *there* (Wei, 1996b, 1997).

2. Assumptions underlying the categorization of morphemes

The basic assumption of the 4-M model is that morphemes are projected differently from the mental lexicon. This paper assumes that differences across abstract lexical entries in the mental lexicon cause different degrees of difficulty in acquiring different types of morphemes. Accordingly, there are three specific assumptions underlying this study.

(1) *Lemmas as abstract entries in the mental lexicon.* Levelt defines “lemma” as the “nonphonological part of an item’s lexical information.” He also states, “lemmas are the driving force behind the speaker’s construction of the surface structure. It is in the lemmas of the mental lexicon that conceptual information is linked to grammatical function” (1989, p. 162). In other words, conceptual information about lexical entries is provided as prelexical feature bundles stored in the mental lexicon as “lemmas.” For example, the prelexical feature bundles of a verb include its semantic-pragmatic selectional restrictions, its argument structure, morphemes for its tense/aspect marking, its exceptional case marking features,

and other necessary information. Under this assumption, lemmas are abstract entries in the mental lexicon that contain information about the three subsystems of lexical structure: “lexical-conceptual structure” conflating universally-available semantic and pragmatic information, “predicate-argument structure” specifying the properties of verbs in terms of their subcategorization frames, how many arguments they may take, and what thematic role each argument receives, and “morphological realization patterns” spelling out surface devices for word order, agreement, tense/aspect marking, and so forth (cf. de Bot & Schreuder, 1993; Myers-Scotton & Jake, 1995).

Lemmas send directions to the “formulator,” which transforms conceptual knowledge of lexical entries into linguistic knowledge in language production (Garrett, 1990; Levelt, 1989; Myers-Scotton, 1993 [1997]). In addition, lemmas can be “directly elected” or “indirectly elected” (Bock & Levelt, 1994), and not all morphemes are activated at the lemma level (Myers-Scotton & Jake, 2000).

(2) *Content versus system morpheme distinction.*¹ Morphemes directly elected at the lemma level are content morphemes (Jake & Myers-Scotton, 1997; Levelt, 1989; Myers-Scotton, 1997). They are content morphemes because “they are ‘directly elected’ by the semantic/pragmatic feature bundle, mapping conceptual structure onto the lemma” (Myers-Scotton & Jake, 2000). The ability to assign or receive a thematic role is a property of content morphemes (Myers-Scotton, 1993 [1997]). Most verbs and some prepositions are prototypical thematic-role assigners, and most nouns are prototypical thematic-role receivers. In contrast, system morphemes neither assign nor receive thematic roles. Inflectional morphemes and most function words are prototypical system morphemes. The content versus system morpheme distinction specified by Myers-Scotton (1993 [1997]) best captures the distinction between thematic and functional elements.² The distinction is important in that it constrains patterns of occurrence of morphemes in bilingual codeswitched speech. The distinction is also important in that it plays a significant role in categorizing morphemes in second language production.

However, different languages may assign particular lexical “concepts” to content or system morphemes differently. Cross-linguistic variation in the assignment of morphemes to either content or system morpheme status is one of the factors affecting IL constructions. This is because target language morphemes may be conceptually congruent with first language counterparts, but they may differ in their status as content or system morphemes (for cross-linguistic variation and analysis, see Jake, 1994, 1998; Myers-Scotton & Jake, 2000).

¹ The content versus system morpheme distinction predicts variation in morpheme membership within particular lexical categories, and this distinction differs from other proposals characterizing the content versus function distinction. Joshi (1985) and Emonds (1985) propose the open versus closed class distinction. This distinction suffers from the deficiencies which assign status on the basis of lexical category membership. For example, not all prepositions are closed class items.

² Abney’s (1987) proposal that the feature [± functional] be category-defining encounters similar deficiencies. For example, while Abney categorizes all pronouns as functional elements, the psycholinguistic (e.g., Friederici, Weissenborn, & Kail, 1991) and codeswitching literature (Myers-Scotton & Jake, 1995; Myers-Scotton & Jake, 2000) provide the evidence that some pronouns in some languages are content morphemes, not system morphemes. For example, in French only emphatic, full form, pronouns such as *moi* ‘me’ and *toi* ‘you’ are content morphemes, and personal pronoun clitics such as *je* ‘I’ and *te* ‘you’ are system morphemes (see Jake, 1994).

Relevant to the study of morpheme accuracy/frequency acquisition are the assumptions underlying the content versus system morpheme distinction. It is assumed that the sources of morphemes refer to differences in the levels of abstract lexical structure. At the lemma level, speaker intentions are mapped onto bundles of semantic/pragmatic features by direct-election (Myers-Scotton & Jake, 1995; Wei, 1996b). This is the mapping of speaker intentions to lemma entries. Also, entries in the mental lexicon support lexical entries. Directly-elected lemmas send directions about lexical-conceptual structure, predicate-argument structure and morphological realization patterns to the formulator; this activates morphosyntactic procedures spelling out the lexical knowledge of the lexical entry. Based on these directions, and after phonological encoding, the formulator produces the surface level string.

(3) *Three types of system morphemes.* The 4-M model specifies three types of system morphemes. As introduced earlier, content morphemes are activated at the lemma level (i.e., the lemmas supporting content morphemes are directly-elected). These directly-elected morphemes may “call” other items required to complete the speaker’s intentions together with content morpheme heads. Such “called” morphemes are also activated at the lemma level, but they are not thematic role assigners or receivers. They are “indirectly-elected” (cf. Bock & Levelt, 1994) system morphemes that are selected from a limited number of choices in the projection of structure required by certain content morphemes. Together with its content morpheme head, an indirectly-elected system morpheme also plays a role in the realization of the predicate-argument structure and morphological realization patterns of the directly-elected morpheme, but it does not represent a lexical concept independent of the directly-elected element with which it is accessed. The 4-M model categorizes indirectly-elected morphemes as “early” system morphemes. In English, for example, the prepositions in *listen to* and *look at* are activated at the lemma level together with the verbs *listen* and *look*.

“Indirect election” plays a role in relating speaker intentions to lemma entries. For example, under certain semantic/pragmatic conditions, with a definite referent, the determiner is indirectly elected by the lemma underlying the noun in English, with a specifying or particularizing effect. For example, the definite article in *the book you gave me yesterday* and the pronoun in *my dog* are indirectly-elected early system morphemes, because the feature of definiteness and the feature of possession are part of the conceptual structure activated by the speaker’s intentions.

In addition to the distinction between direct-election and indirect-election at the lemma level, the 4-M model categorizes another type of system morphemes as “late” system morphemes. They are neither activated at the lemma level, nor are they thematic role assigners or receivers. The 4-M model further subcategorizes late system morphemes into two types: “bridges” and “outsiders.”

Unlike an early system morpheme, a late “bridge” system morpheme is not activated at the conceptual level with a content morpheme, but rather integrates a content morpheme into a larger constituent. For example, the genitive/possessive *of* in English as in *the property of the church* is a late bridge system morpheme. The preposition *of* is structurally required in the English grammar to integrate the [NP] *the church* into the [NP] *the property* in order to realize the possessive case. Similarly, while the preposition *of* in *I’ve never thought of that* is an early system morpheme (i.e., indirectly-elected), the preposition *of* in *students of linguistics* is a late bridge system morpheme.

Unlike a late bridge system morpheme, which “depends on information within the maximal projection in which it occurs”, a late “outsider” system morpheme “depends on grammatical information outside of their (its) own maximal projection” (Myers-Scotton & Jake, 2000). In other words, late outsider system morphemes are structurally assigned at the positional/surface level (i.e., required in the “spelling out” of aspects of the morphological realization patterns). For example, 3rd person present tense *-s* and auxiliary verbs like *do*, *have*, *be*, and *will/shall* (future) are late outsider system morphemes, because they are structurally assigned by the English grammar.

The 4-M model includes as late outsider system morphemes any morpheme under INFL which cannot be realized without coindexing with another element in the sentence. Accordingly, tense marking in English also belongs to this category. Thus, the distinction between late bridge and late outsider system morphemes should be clear: the former referring not to the head of the maximal projection in which they occur, but to other information in this maximal projection, and the latter referring to information outside their maximal projection of the head.

In conclusion, the 4-M model categorizes morphemes into four types based on how they are elected. While information about all types of morphemes is present in lemmas, information about content morphemes and early system morphemes is salient at the conceptual level, and information about late system morphemes becomes salient at the positional level of the formulator.

3. Variation within individual lexical categories

Although members of specific morpheme categories are typically content or system morphemes, there is variation within a particular language regarding category membership of specific lexical items (for cross-linguistic differences in morpheme classification, see Jake, 1994, 1998; Myers-Scotton & Jake, 2000). While it is true that in English most nouns and verbs are content morphemes and that most elements in INFL are system morphemes, the content versus system morpheme distinction is not a lexical category-defining feature. For example, items projected under INFL can be content or system morphemes. While auxiliary verbs such as *do*, *have*, *be* and *will/shall* (future) are system morphemes for tense/aspect marking, mapping predicate-argument structure onto morphological realization patterns, modal verbs such as *can*, *should*, *may* and *must* are content morphemes, conveying lexical-conceptual structure (cf. Dowty, 1979; Levelt, 1989; Myers-Scotton & Jake, 1995; Wei, 1996b, 1997). Below are some of the most important examples of variation within specific lexical categories. Some of these categories are studied here.

Within the category of determiners, some members are early system morphemes, such as possessive *my*, *your*, *their*, and so forth and demonstrative *this/these* and *that/those*, because the semantic and pragmatic features that activate the content morphemes they occur with point to and require the definite and deictic features to be “fleshed out.” Similarly, definite and indefinite articles can be early system morphemes if they are conceptually activated by the nouns with which they occur. For example, the articles in *I visited **the** new library again with **a** new student* are early system morphemes. However, some determiners are late system morphemes because they are structurally assigned by the grammar. For example, the articles in *in **the** hospital* (American English), *in **the** future* (American English), *in **a** bind*, *all of **a** sudden* are late bridge system morphemes.

Within the category of pronouns, some members are content morphemes because they receive thematic roles and occur in grammatical argument positions, such as subject or object. For example, personal pronouns such as *I/me*, *they/them*, *he/him*, and so forth and freestanding demonstratives like *this/these* and *that/those* (e.g., *this is the right answer*) are all content morphemes. Others are structurally-assigned (late) system morphemes. For example, dummy pronouns *it* and *there* do not receive thematic roles and only serve to satisfy surface subject requirements of the grammar of English (Jake, 1994; Wei, 1996b, 1997).

Within the category of prepositions, some members are content morphemes because they assign thematic roles, as well as case. For example, the preposition *for* assigns the thematic role of beneficiary or goal to *Mary* in *Steve bought a Spanish-English dictionary for Mary*. Others are system morphemes. For example, the prepositional satellites in verb phrases such as *listen to the radio*, *look at the picture*, *depart from the normal procedure*, *rely on his assistance*, and *think of a better solution* are early system morphemes. This is because they are “called” by their respective verbs, which are content morphemes activated at the lemma level. Furthermore, other prepositions are late system morphemes because they only assign case. For example, the preposition *of* in *students of linguistics* or in *friends of my family* is a late system morpheme.

Thus, there is variation in category membership within each individual lexical category, and variation exists in the fact that morphemes are projected differently from the mental lexicon: conceptually or nonconceptually. Also, early and late system morphemes differ in that the former include semantic and pragmatic features salient at the lexical-conceptual level, whereas the latter do not.

4. Evidence for the 4-M model

Reported here are some of the findings relevant to testing the categorization of morphemes as specified in the 4-M model. The data come from a comprehensive research project on adult second language morpheme acquisition (Wei, 1996b), which studied learners’ accuracy/frequency acquisition of English inflectional morphemes, auxiliary verbs, existential verbs, determiners, pronouns, prepositions, and IL constructions driven by the morphemes under investigation.

The original study had 60 subjects who were native speakers of Chinese and Japanese learning English as a second language. They were divided into three groups for each L1 background based on their English proficiency levels: prebasic, basic, or beyond-basic, following the learner stages identified in Klein and Perdue (1993, pp. 30–40) and Klein, Dietrich and Noyau (1993, pp. 104–110).

The data were collected from designed tasks (cf. Hatch & Farhady, 1982; Larsen-Freeman & Long, 1991); the interview consisted of two parts: (1) a series of questions designed to approximate natural conversation; (2) two picture description tasks, one designed to elicit descriptions related to “existence, location, possession, condition, etc.” and the other designed to elicit descriptions involving “ongoing, completed, or future activities.” Each interview was tape-recorded and coded for the specific morphosyntactic categories and subcategories under investigation.

Error identification and frequency count followed three procedures. (1) “Obligatory

occasion analysis" (Ellis, 1994, p. 91) was employed to count the frequency of occurrence of the morphemes under investigation. In a particular verbal interaction context, the learner was obliged to produce a target linguistic item. If it was not produced or produced incorrectly, it was counted as an error. The learner's "avoidance" of the production of a certain item was not counted as "missing," provided that the sentence itself was grammatically acceptable. (2) Any linguistic item relevant to the investigation produced in free production (picture description tasks) was judged according to individual morphosyntactic rules. (3) The acceptability of a linguistic item under investigation must pass the "filter" of three native speakers' judgments (three American professors of linguistics).

The statistical test adopted for the study was Poisson Regression, which models frequency of occurrence and predicts performance on the dependent variables via one or more independent variables (cf. Hatch & Farhady, 1982). Poisson Regression expands the possible sources of prediction and tests to see which of the many variables, and which combination of variables, allow us to make the best prediction of frequency of occurrence. Statistical significance of difference between variables is set at $Z > 3.16$ in order to reject the null hypothesis (Finch, 1996).

The study proposed an implicational acquisition order based on how different types of morphemes are activated. Accordingly, two hypotheses were tested:

¥Directly-elected content morphemes are produced accurately before system morphemes.

¥Conceptual "salience" is a significant factor in system morpheme acquisition: early system morphemes are produced before late ones.

These two hypotheses result in a generalized acquisition ***Implicational Hierarchy Principle***:

directly-elected > indirectly-elected > structurally-assigned
(content morphemes) (early system morphemes) (late system morphemes).

(">" = to be acquired before ...)

5. Results of the tasks

Reported here are the data from the prebasic and basic stage Chinese and Japanese learners' accuracy/frequency production of English determiners and pronouns.

5.1 Determiners

Table 1 overleaf shows that both the prebasic stage Chinese and Japanese learners have extreme difficulty in correctly producing both early and late system articles. However, the same learners do not have serious difficulty with the early system possessives and demonstratives. The different degrees of learning difficulty with the early and late system articles are clearly shown in the basic stage. Table 2 also overleaf shows that the Z scores for these two types of articles indicate that their difference in the basic stage are statistically significant.

The types of errors produced by these learners are illustrated below.

Table 1
English determiners: frequency/% correct production in obligatory contexts

Morpheme types:	Early system morphemes						Late system morphemes	
Variables:	possessive		demonstrative		article (E)		article (L)	
Prebasic stage								
Chinese	112/136	82%	56/67	84%	12/125	10%	1/33	3%
Japanese	96/113	85%	64/67	96%	10/110	9%	6/50	12%
Basic stage								
Chinese	184/191	96%	104/105	99%	137/214	64%	8/43	19%
Japanese	126/139	91%	87/87	100%	96/130	74%	11/40	28%

E: early system morpheme L: late system morpheme

Table 2
Statistical test results

Variables:	art. (L) versus poss.	art. (E) versus poss.	art. (L) versus dem.	art. (E) versus dem.	art. (L)/art. (E)
Prebasic Stage					
Chinese	*Z=5.68	*Z=6.75	*Z=6.71	*Z=7.53	NS
Japanese	*Z=4.52	*Z=7.14	*Z=5.14	*Z=7.38	NS
Basic stage					
Chinese	*Z=3.30	*Z=6.15	*Z=3.27	*Z=5.60	*Z=6.12
Japanese	*Z=9.40	*Z=4.76	*Z=10.05	*Z=6.09	*Z=6.76

Z>3.16

*: significant; NS: not significant

Examples of errors with determiners: ([Ø]=missing)

Chinese L1

Article (L):

- (1) *I study all [Ø] time, just study English.* [TL: all **the** time]
- (2) *I not listen to radio.* [TL: listen to **the** radio]

Articles (E):

- (3) (Describing a picture) *[Ø] Man eating [Ø] hot dog is her father ... maybe.* [TL: **the** man ... **a** hot dog ...]
- (4) *My parents have **the** three children.* [TL: ... have three children]

Possessive:

- (5) (Interviewer: Does your daughter go to school here?) *Yes, [Ø] daughter go to school ... yes, [Ø] daughter go to school.* [TL: **my/our** daughter]
- (6) (Describing a picture) *I don't know why **this** people like sit on beach in summer.* [TL: **these/those** people]

Japanese L1

Article (L):

(7) *When in Japan I not know gun control, but now in America all of [Ø] sudden I care about gun control.* [TL: all of **a** sudden](8) *I usually meet my friends at [Ø] weekend.* [TL: at **the** weekend]

Article (E):

(9) *I not like [Ø] hot weather in South Carolina here.* [TL: **the** hot weather](10) *I have [Ø] brother in Japan.* [TL: **a** brother]

Possessive:

(11) *[Ø] Husband come here first. I and [Ø] daughter come after one year.* [TL: **my** husband ... **my** daughter]

Demonstrative:

(12) (Describing a picture, pointing to the boys and girls playing on the sand and then to the couple sitting nearby) *I think [Ø] children are [Ø] couple's kids.* [TL: **these/those** children ... **this/that** couple's kids]

5.2 Pronouns

Table 3 overleaf shows that both the prebasic stage Chinese and Japanese learners have far more difficulty in correctly producing the late system pronouns (dummy pronouns *it* and *there*) than in correctly producing the content pronouns (personal and demonstrative). In contrast, although both the basic stage Chinese and Japanese learners still have some difficulty with the late system pronouns, they show very clear progress in correctly producing them. Table 4 also overleaf shows that the Z scores for these two types of pronouns indicate that their differences are statistically significant in the prebasic stage learning.

The types of errors produced by these learners are illustrated below.

Examples of errors with pronouns: ([Ø] = missing)Chinese L1Dummy *it*:(13) *In summer [Ø] not hot.* [TL: **it** is not hot](14) *[Ø] Very difficult find job.* [TL: **it** is very difficult ...]Dummy *there*:(15) *In China [Ø] too many people.* [TL: **there** are ...](16) *I think [Ø] only seven students in that ESL class.* [TL: **there** are ...]

Personal:

(17) (Interviewer: Does your family support you?) *Yes, yes, my family always support [Ø] sometimes give [Ø] money.* [TL: ... support **me**..... give **me** money]

Demonstrative:

(18) (Pointing to the pictures on the table) *Pictures here? Oh, **this** I think are beautiful pictures.* [TL: **these/those**]

Table 3
English pronouns: frequency/percent correct production in obligatory contexts

Morpheme types:	Content morphemes				Late system morphemes			
Variables:	personal		demonstrative		it		there	
Prebasic Stage								
Chinese	421/458	92%	35/38	92%	15/48	31%	4/48	8%
Japanese	536/539	94%	32/34	94%	17/37	46%	13/40	32%
Basic Stage								
Chinese	841/852	99%	36/37	97%	39/62	63%	26/31	84%
Japanese	627/635	99%	46/47	98%	52/63	83%	30/34	88%

Table 4
Statistical test results

Variables:	it vs. dem.	there vs. dem.	it vs. person	there vs. person
Prebasic Stage				
Chinese	*Z=7.14	*Z=3.64	*Z=6.53	*Z=3.76
Japanese	*Z=3.35	*Z=3.76	*Z=3.85	*Z=4.56

Z >3.16
*: significant

Japanese L1

Dummy *it*:

- (19) *But in Japan [Ø] impossible live in big apartment.* [TL: **it** is impossible ...]
(20) (Describing a picture) *On beach [Ø] very hot.* [TL: **it** is very hot]

Dummy *there*:

- (21) *In Columbia [Ø] really kindness or something, I mean.* [TL: **there** is ...]
(22) *We now live in apartment. [Ø] Two bedroom and one bathroom in apartment.* [TL: **There** are ...]

Personal:

- (23) (Interviewer: Did you get the application form from the EPI office?) *No, no, my friend gave [Ø] to me.* [TL: ... give **it** to me]

Demonstrative:

- (24) (Interviewer: Do you think many Japanese students want to do this? (Talking about studying English in America)) *Some Japanese students ... some young students want to do **these**, not all students want to do **these**.* [TL: ... want to do **this/that**]

6. Major findings

The above accuracy/frequency count of the learners' acquisition of the morpheme types reported in this paper provides statistical evidence in support of the 4-M model and the hypotheses formulated for the study. Cross-linguistically, some meaningful patterns of morpheme accuracy/frequency acquisition have emerged.

1. Not all determiners are produced with equal accuracy in SLA. Learner errors with the late system articles (*a/the*) present constant difficulty to both the prebasic and basic stage Chinese and Japanese learners. The basic stage learners show far more progress with the early system articles (*a/the*). Distinguishing the way that articles are elected or accessed explains the otherwise seemingly random variation in these data. As predicted by the ***Implicational Hierarchy Principle***, the same learners in both stages do not show much difficulty in producing accurately the early system determiners (possessive and demonstrative).

Results also show that although the early system article is an indirectly-elected determiner, it shows more statistically significant learning difficulty than the other indirectly-elected determiners. The possible reason for this difference is that possessive *my/your/ his*, and so forth and demonstrative *this/these* and *that/those* may carry more semantic weight than the indirectly-elected article *a/the*. Within the same morpheme category, the more “meaningful” morphemes are acquired before the less “meaningful” ones. However, the linguistic behavior of the early system articles needs more study.³

³ Potential influence from the learners’ L1 may exist in the fact that neither Chinese nor Japanese possesses articles to express their equivalent meanings in English. In these languages, “definiteness” is expressed by specific early system morphemes such as possessives like *my, your, her, its*, and so forth or demonstratives like *this, that, these, and those*. Consequently, the prebasic stage Chinese and Japanese learners may fail to produce the English definite article if the “definiteness” of a referent is less obvious than that indicated by possessives or demonstratives in their L1. For example, the Chinese sentence below is totally grammatical.

wo zuotian canguan le bowuguan.
I yesterday visit PART/PERF museum
‘Yesterday I visited the museum.’

Similarly, the Japanese sentence below is totally grammatical.

kyooshitsu ni Mimura san wa imasen.
classroom PREP/in Mimura Mr. PART/TOP be not
‘Mr. Mimura is not in the classroom.’

The English indefinite article may also cause learning difficulty to the prebasic stage Chinese and Japanese learners. This is because neither Chinese nor Japanese possesses the similar early system morpheme as the English indefinite article *a/an*. In these languages, the realization of “indefiniteness” with the meaning of ‘one’ or ‘anyone’ does not require a specific early system morpheme. For example, the Chinese sentence below is totally grammatical.

ni youmeiyou che?
you have-not-have car
‘Do you have a car?’

Similarly, the Japanese sentence below is totally grammatical.

kare wa hon o yon-de iru.
he PART/TOP book PART/OBJ read-PROG AUX
‘He is reading a book.’

The similar potential influence from the learners’ L1 may also occur when the prebasic stage Chinese and Japanese learners produce English verbs with satellites. As observed, English early system morphemes such as prepositions *to* in *listen to*, *at* in *look at*, and *for* in *wait for*, and particles *up* in *pick up*, *off* in *take off*, and *on* in *put on* are often missed in the prebasic stage learners’ speech production. This is because neither Chinese nor Japanese requires the similar satellites to complete the specific meanings of the verbs in question. Consequently, these learners may encounter some difficulties in producing certain English verbs without paying attention to the early system morphemes as required by these verbs.

2. Not all pronouns are produced with equal accuracy. Learner errors with the late system pronouns (dummy pronouns *it* and *there*) persistently occur. Both the prebasic Chinese and Japanese learners have almost no difficulty in acquiring the content pronouns (personal *I/me, he/him*, etc. and demonstrative *this/these, that/those*), but have serious difficulty in producing accurately the late system pronouns. Even the basic stage learners still show more difficulty with the late system pronouns than with the content ones.

7. Implications of the study

The predictions based on the generalized acquisition *Implicational Hierarchy Principle* have been supported. Under the assumption that accurate production reflects acquisition, one can claim that directly-elected morphemes (content morphemes) are acquired before system morphemes, and indirectly-elected morphemes (early system morphemes) are acquired before structurally-assigned ones (late system morphemes). There are three major findings regarding second language morpheme acquisition.

1. Cross-linguistically, the distinctions among the types of morphemes operate in second language morpheme acquisition order. Both the prebasic and basic stage learners experience more serious learning difficulties with late system morphemes than with early system and content morphemes. Although the prebasic stage learners may experience a similar degree of learning difficulties with both late and early system morphemes within certain morphosyntactic categories, the basic stage learners show more progress in acquiring early system morphemes than late ones.
2. The distinctions among the types of morphemes have consequences for the form of IL. Target language (TL) late system morphemes must be acquired before TL morphological realization patterns become possible (Wei, 1996a). Incomplete or nonacquisition of TL late system morphemes results in nontarget-like grammatical constructions or ungrammaticality. This is shown in the learner error examples like (5) *Yes, daughter go to school ... yes, daughter go to school*. TL early system morphemes must be accessed together with their content morpheme heads before TL predicate-argument structures become native-like (cf. “form before function” in Perdue, 1993).
3. It is the different projections of morphemes from the mental lexicon that decide the degrees of learning difficulty (Bock & Levelt, 1994; Jake & Myers-Scotton, 1997; Levelt, 1989; Myers-Scotton & Jake, 2000; Wei, 1996b). It is the sources of morphemes that can effectively explain why certain learner errors are more frequent and common than others and thus determine the order of second language morpheme accuracy/frequency acquisition.

8. Conclusion

This study has provided evidence from IL data for the 4-M model of morpheme activation. This evidence has three major implications for SLA research.

⁴ White (1989, 1991) assumes that UG will not be involved in the acquisition of a large part of the lexicon and “properties that are specific to a language will have to be learned.” According to her, “words and their meanings will have to be learned, together with their syntactic categories and subcategorization requirements” (1989, p. 30).

1. Lexical categories are language-specific and must be learned as such (White, 1989, 1991).⁴ Sufficient acquisition of TL lexical categories, including their subcategories (i.e., variation within specific morphosyntactic categories), must be acquired in order to set or reset TL parameters. That is, the acquired TL lexical structure will eventually replace the L1 lexical structure. The major differences between L1 and L2 acquisition lie in their respective mechanisms and processes, rather than in the abstract universal categories of language.
2. SLA processes and developmental patterns can be best explained and predicted in terms of the nature of different types of morphemes being acquired. The differential projection of morphemes decides the acquisition order. Directly-elected morphemes (content morphemes) are acquired first and they add more lexical specifications, resulting in the projection of indirectly-elected morphemes (early system morphemes). The acquisition of TL structurally-assigned morphemes (late system morphemes) will result in the realization of more target-like morphological realization patterns.
3. While commonly observed IL variability in terms of morphosyntactic structures exists, most of this reflects the nonacquisition or non-native-like acquisition of TL structurally-assigned system morphemes, rather than L1 influence. L1 influence seems to be of two main types: Indirectly-elected system morphemes may fail to be accessed together with their content morphemes (e.g., as observed, prebasic stage Chinese and Japanese learners may leave out morphemes such as articles *the* in *we can use **the** same book* and *a* in *I have **a** brother in Japan.*); L1 lexical-conceptual structure (i.e., semantics) may be substituted for L2 lexical-conceptual structure in L2 lexical items (e.g., as observed, early stage Chinese learners may say *house* for *apartment*, *red tea* for *black tea*, *sit in the bus* for *take the bus*, etc.).

In conclusion, the 4-M model categorizes morphemes into four types in terms of how they are accessed in language production: at the lemma level or at the level of the formulator. Content morphemes are activated at the lemma level, where the speaker's semantic and pragmatic intentions are expressed. What is new and crucial in the 4-M model is the proposal that system morphemes be further categorized into two classes in terms of how they are accessed: early and late system morphemes. Early system morphemes are also activated at the lemma level together with their content morpheme heads for their maximal projection. Late system morphemes are subcategorized into bridge and outsider morphemes. Unlike content morphemes and early system morphemes, late system morphemes depend on other types of information for their activation, and this information is only available at the level of the formulator, where language specific morphosyntactic patterns must be realized.

This study offers some concrete evidence for the 4-M model from data dealing with second language morpheme accuracy/frequency acquisition. The hypotheses formulated according to this model are confirmed: within and across the morpheme types, content morphemes are acquired before system morphemes, and early system morphemes are acquired before late ones. The study concludes that IL constructions are driven by the unequal acquisition of different types of morphemes. Second language morpheme acquisition order is determined by the sources of morphemes, and IL development is a process of acquiring different types of TL morphemes at a different rate. The 4-M model

provides a transparent window through which the nature of different types of morphemes in relation to morpheme acquisition order and its resulting IL constructions can be meaningfully described and insightfully explained.

References

- ABNEY, S. P. (1987). The English noun phrase in its sentential aspect. *MIT Working Papers in Linguistics*. Cambridge, MA: MIT Press.
- BAILEY, N., MADDEN, C., & KRASHEN, S. D. (1974). Is there a "natural sequence" in adult second language learning? *Language Learning*, **24**, 235–244.
- BOCK, K., & LEVELT, W. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). New York: Academic Press.
- De BOT, K., & SCHREUDER, R. (1993). Word production and the bilingual lexicon. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 191–214). Amsterdam: John Benjamins.
- DOWTY, D. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht: D. Reidel.
- DULAY, H. C., & BURT, M. K. (1972). Goofing: An indicator of children's second language strategies. *Language Learning*, **22**, 234–252.
- DULAY, H. C., & BURT, M. K. (1973). Should we teach children syntax? *Language Learning*, **23**, 245–258.
- DULAY, H. C., & BURT, M. K. (1974). Natural sequences in child second language acquisition. *Language Learning*, **24**, 37–53.
- DULAY, H. C., BURT, M. K., & KRASHEN, S. D. (1982). *Language two*. Rowley, MA: Newbury House.
- ELLIS, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- EMONDS, J. (1985). *A unified theory of syntactic categories*. Dordrecht: Foris.
- FINCH, H. (1996). Personal communication. The Statistics Laboratory at the University of South Carolina, Columbia, SC.
- FRIEDERICI, A. D., WEISSENBORN, J., & KAIL, M. (1991). Pronoun comprehension in aphasia: A comparison of three languages. *Brain and Language*, **41**, 289–310.
- GARRETT, M. F. (1990). Sentence processing. In D. N. Osherson & H. Lasnik (Eds.), *An invitation to cognitive science, Vol. 1* (pp. 133–175). Cambridge, MA: MIT Press.
- HATCH, E., & FARHADY, H. (1982). *Research design and statistics for applied linguistics*. Rowley, MA: Newbury House.
- JAKE, J. L. (1994). Intrasentential codeswitching and pronouns: On the categorial status of functional elements. *Linguistics*, **32**, 271–298.
- JAKE, J. L. (1998). Constructing interlanguage: Building a composite matrix language. *Linguistics*, **36**, 333–382.
- JAKE, J. L., & MYERS-SCOTTON, C. (1997). Relating interlanguage to codeswitching: The composite matrix language. *Proceedings of Boston University Conference on Language Development*, **21**, 319–330. Boston, MA: Cascadia Press.
- JOSHI, A. (1985). Processing of sentences with intrasentential codeswitching. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing: Psychological computation and theoretical perspectives* (pp. 190–205). Cambridge, U.K.: Cambridge University Press.
- KLEIN, W., DIETRICH, R., & NOYAU, C. (1993). The acquisition of temporality. In C. Perdue (Ed.), *Adult language acquisition: Cross-linguistic perspectives, Vol. 2* (pp. 73–118). Cambridge, U.K.: Cambridge University Press.
- KLEIN, W., & PERDUE, C. (1993). Utterance structure. In C. Perdue (Ed.), *Adult language acquisition: Cross-linguistic perspectives, Vol. 2* (pp. 3–40). Cambridge, U.K.: Cambridge University Press.

-
- LARSEN-FREEMAN, D., & LONG, M. H. (1991). *An introduction to second language acquisition research*. London: Longman.
- LEVELT, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- MYERS-SCOTTON, C. (1993 [1997]). *Duelling languages: Grammatical structure in codeswitching*. Oxford: Clarendon Press. 1997 paperback edition with new "Afterword."
- MYERS-SCOTTON, C. (1997). Codeswitching. In Florian Coulmas (Ed.), *Handbook of sociolinguistics* (pp.217–237). Oxford: Blackwell.
- MYERS-SCOTTON, C., & JAKE, J. L. (1995). Matching lemmas in a bilingual competence and performance model: Evidence from intrasentential codeswitching. *Linguistics*, **33**, 981–1024.
- MYERS-SCOTTON, C., & JAKE, J. L. (2000). Explaining aspects of codeswitching and their implications. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 91–125). Oxford: Blackwell.
- PERDUE, C. (Ed.). (1993). *Adult language acquisition: Cross-linguistic perspectives*, Vol. 2. Cambridge, U.K.: Cambridge University Press.
- WEI, L. (1996a). Organizational principles behind codeswitching and interlanguage development in early adult second language acquisition. In J. Arnold, R. Blake, B. Davidson, S. Schwenter & J. Solomon (Eds.), *Sociolinguistic variation: Data, theory, and analysis. Selected papers from NWAV 23* (pp.417–431). Stanford, CA: Center for the Study of Language and Information.
- WEI, L. (1996b). *Variation in the acquisition of morpheme types in the interlanguage of Chinese and Japanese learners of English as a second language*. Unpublished Ph.D. dissertation, Linguistics Program, University of South Carolina, Columbia, SC.
- WEI, L. (1997). *Variation in morpheme acquisition and its effects on interlanguage constructions*. Paper presented at Second Language Research Forum (SLRF 1997), Michigan State University, East Lansing, MI.
- WHITE, L. (1989). *Universal grammar and second language acquisition*. Amsterdam: John Benjamins.
- WHITE, L. (1991). Argument structure in second language acquisition. *Journal of French Language Studies*, **1**, 189–207.
-